Math 115B Running Notes

Brendan Connelly

January to March 2025

Dual Spaces

Proposition. For any two vector spaces V and W, the set of linear functions from V to W, denoted $\mathcal{L}(V, W)$, is a vector space with the following operations:

- **Addition:** For $S, T \in \mathcal{L}(V, W)$, define $(S + T)(v) = S(v) + T(v)$ for all $v \in V$.
- **Scalar Multiplication:** For $\alpha \in k$ and $T \in \mathcal{L}(V, W)$, define $(\alpha T)(v) = \alpha(T(v))$ for all $v \in V$.

Definition (Dual Vector Space)**.** For any vector space V, the *dual vector space* V^* is the set of all linear functions from V to k, denoted:

$$V^* := \mathcal{L}(V, k).$$

Definition. Given a vector space V, the elements of the dual vector space V^* are known as *linear functionals*.

Proposition. For any basis $\beta = \{v_1, \ldots, v_d\}$ of a finite-dimensional vector space V, there exists an isomorphism

$$[-]_\beta : \mathcal{L}(V, V) \to k^{d \times d}$$

defined by the formula:

$$[T]_\beta = \begin{pmatrix} [T(v_1)]_\beta & [T(v_2)]_\beta & \cdots & [T(v_d)]_\beta \end{pmatrix},$$

for any $T \in \mathcal{L}(V, V)$.

Theorem (2.20)**.** Let V and W be finite-dimensional vector spaces over $\mathbb{K}$, and let $\beta = \{v_1, \ldots, v_m\}$ be a basis for V, and $\gamma = \{w_1, \ldots, w_n\}$ be a basis for W. Then there exists a linear isomorphism:

$$[-]_{\gamma,\beta} : \mathcal{L}(V, W) \to k^{n \times m}.$$

Corollary. If V is a vector space of dimension m and W is a vector space of dimension n, then:

$$\dim(\mathcal{L}(V, W)) = mn.$$

Corollary. If V is a finite-dimensional vector space, then:

$$\dim(V^*) = \dim(V).$$

Definition (Dual Basis Vector)**.** Given a finite-dimensional vector space V and a basis $\beta = \{v_1, \dots, v_d\}$ of V, the i-th *dual basis vector* is the linear functional $v_i^* : V \to k$ defined by the formula:

$$v_i^*(\vec{v}) = \alpha_i,$$

where $\vec{v} = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_d v_d$ is the representation of $\vec{v} \in V$ in terms of the basis β.

Theorem (2.24)**.** If V is a finite-dimensional vector space and $\beta = \{v_1, \dots, v_d\}$ is a basis for V, then the set $\{v_1^*, v_2^*, \dots, v_d^*\}$ is a basis for V^*. Moreover, for any $f \in V^*$, we have:

$$f = f(v_1)v_1^* + f(v_2)v_2^* + \cdots + f(v_d)v_d^*.$$

Definition (Dual Basis)**.** If V is a finite-dimensional vector space with basis $\beta = \{v_1, \dots, v_d\}$, the basis

$$\beta^* = \{v_1^*, \dots, v_d^*\}$$

is called the *dual basis.*

Theorem (2.25)**.** Let V, W be finite-dimensional vector spaces over k. Let $\mathcal{B}$ be a basis for V and $\mathcal{C}$ be a basis for W. Let $T : V \to W$ be a linear transformation. Then, $T^* : W^* \to V^*$, given by

$$T^*(g) = g \circ T \quad \text{for any } g \in W^*,$$

is linear. Moreover,

$$[T^*]_{\mathcal{C}^*}^{\mathcal{B}^*} = \left([T]_{\mathcal{B}}^{\mathcal{C}}\right)^t.$$

Theorem (2.26)**.** Let V be a finite-dimensional vector space. Then, the map

$$\Psi : V \to (V^*)^*$$

given by the formula

$$\Psi(v)(f) = f(v),$$

for $v \in V$ and $f \in V^*$, is a linear isomorphism.

Remark (On Dual and Double Dual Vector Spaces)**.**

$$\dim(V) = \dim(V^*) = \dim((V^*)^*),$$

We also know that $V^* = \mathcal{L}(V, k)$, so

$$(V^*)^* = \mathcal{L}(V^*, k).$$

The main point is that for any finite-dimensional vector space V, there exists an isomorphism between V and its double dual $(V^*)^*$. This isomorphism does not depend on the choice of a basis.

Remark. *First equality easily shown in hw. Second equality is easy for one inclusion*

$$\ker(T^*) = (\operatorname{Im} T)^\circ.$$

$$\operatorname{Im}(T^*) = (\ker T)^\circ.$$

If W is a subspace of V where V may actually be infinite dimensional, then

$$\dim(W) + \dim(W^0) = \dim(V).$$

1 Eigenvalues, Eigenvectors, & Diagonalizability

Definition (Eigenvector/Eigenvalue)**.** Assume $T : V \to V$ is linear, where V is a vector space. We say $v \in V$ is an eigenvector of T with eigenvalue $\lambda \in k$ if

$$T(v) = \lambda v \quad \text{and} \quad v \neq 0.$$

Definition (Diagonalizable)**.** A linear transformation $T : V \to V$, where V is a finite-dimensional vector space, is said to be *diagonalizable* if there exists a basis $\mathcal{B}$ of V such that the matrix $[T]_\mathcal{B}$ is a diagonal matrix.

Theorem (5.1)**.** Let V be a finite-dimensional vector space and $T : V \to V$ a linear transformation. Then, T is diagonalizable if and only if there exists a basis $\mathcal{B} = \{v_1, \ldots, v_d\}$ for V such that for any $i \in \{1, 2, \ldots, d\}$, v_i is an eigenvector of T with some eigenvalue $\lambda_i \in k$.

Theorem (5.2)**.** T has $\lambda \in k$ as an eigenvalue if and only if $\ker(T - \lambda I) \neq \{0\}$.

Corollary. λ is an eigenvalue of T if and only if

$$\det(T - \lambda I) = 0.$$

Definition (Determinant)**.** The determinant $\det(A) \in k$ is defined as given in the textbook on page 205.

Definition (Characteristic Polynomial)**.** The characteristic polynomial of $A \in k^{n\times n}$ is

$$\det(T - \lambda I) \in k[\lambda].$$

Definition (Determinant)**.** The determinant of a linear endomorphism $T : V \to V$ of a finite-dimensional vector space V is defined as

$$\det([T]_\mathcal{B}),$$

where $\mathcal{B}$ is a basis for V and $[T]_\mathcal{B}$ is the matrix representation of T with respect to $\mathcal{B}$.

Theorem (5.3)**.** The characteristic polynomial of T is a polynomial of degree n, where $n = \dim(V)$, and the coefficient on t^n is 1. More precisely, it is $(-1)^n$.

Corollary (Number of eigenvalues)**.** Because any polynomial $P_n(\lambda)$ can have at most n-roots (over any field), we conclude:

$$\text{If } \dim(V) = n, \text{ then } T \text{ has at most } n \text{ eigenvalues.}$$

Definition (Polynomial Splits)**.** A polynomial $p(t) \in k[t]$ splits over k if there exist $c, a_1, \ldots, a_d \in k$ such that

$$p(t) = c(t - a_1) \cdots (t - a_d).$$

Theorem (5.6)**.** [Diagonalizability and Splitting] If T is diagonalizable, then the characteristic polynomial of T splits over k.

Rmk: This is only a one-way implication. You can use the contrapositive to show that T is not diagonalizable.

Definition (Algebraic Multiplicity)**.** Given an eigenvalue λ of T, the *algebraic multiplicity* of λ is the largest positive integer j such that $(t - \lambda)^j$ divides the characteristic polynomial of T.

Definition (Eigenspace)**.** Given an eigenvalue λ of T, the *eigenspace* for λ is the span of its eigenvectors with eigenvalue λ. Denote this eigenspace by V_λ.

Example: $V_\lambda = \text{span}\{\text{eigenvectors of } T \text{ with eigenvalue } \lambda\}$.

Definition (Geometric Multiplicity)**.** Given an eigenvalue λ of T, its *geometric multiplicity* is $\dim(V_\lambda)$.

Theorem (5.7)**.** If λ is an eigenvalue for T and has algebraic multiplicity m, then

$$\dim(V_\lambda) \leq m.$$

Equivalently,

$$\text{geo}(\lambda) \leq \text{alg}(\lambda).$$

Theorem (5.8)**.** T is diagonalizable if and only if for every eigenvalue λ_i of T, the geometric multiplicity of λ_i equals its algebraic multiplicity:

$$\text{geo}(\lambda_i) = \text{alg}(\lambda_i).$$

2 Cayley-Hamilton

Theorem (Cayley-Hamilton)**.** If $A \in k^{n \times n}$ and the characteristic polynomial of A is

$$(-1)^d t^d + a_{d-1} t^{d-1} + \cdots + a_1 t + a_0,$$

where $a_{d-1}, \ldots, a_0 \in k$, then

$$(-1)^d A^d + a_{d-1} A^{d-1} + \cdots + a_1 A + a_0 I = 0,$$

where 0 is the zero matrix.

Definition (Nilpotent Maps/Matrices)**.** T is nilpotent if $T^k = 0$ for some $k \in \mathbb{N}$.

Proposition (Eigenvalues of Nilpotent Matrices)**.** Let T be a nilpotent linear map (or matrix). Then, the only eigenvalue of T is 0.

Proof. Suppose T is nilpotent, so there exists some positive integer k such that $T^k = 0$.

Let λ be an eigenvalue of T with corresponding eigenvector $v \neq 0$, i.e.,

$$T(v) = \lambda v.$$

Applying T^k to v, we get:

$$T^k(v) = T^{k-1}(T(v)) = T^{k-1}(\lambda v) = \lambda T^{k-1}(v).$$

Repeating this process iteratively, we find:

$$T^k(v) = \lambda^k v.$$

However, since $T^k = 0$, it follows that:

$$T^k(v) = 0 = \lambda^k v.$$

Because $v \neq 0$, we must have $\lambda^k = 0$. The only solution in the field of scalars (typically $\mathbb{C}$ or $\mathbb{R}$) is $\lambda = 0$. Therefore, the only eigenvalue of a nilpotent matrix T is 0. □

Corollary (Cayley-Hamilton for Linear Transformations)**.** Let $T : V \to V$ be a linear transformation for V a finite-dimensional vector space over a field k, and let

$$p(t) = (-1)^{\dim(V)} t^d + a_{d-1} t^{d-1} + \cdots + a_1 t + a_0$$

be the characteristic polynomial for T.

Then, in $\mathcal{L}(V, V)$,

$$p(T) = (-1)^{\dim(V)} T^d + a_{d-1} T^{d-1} + \cdots + a_1 T + a_0 I = 0.$$

Definition (T-invariant Subspace)**.** A subspace W of V is called T*-invariant* if $T(W) \subseteq W$, i.e.,

$$\{T(w) \mid w \in W\} \subseteq W.$$

Proposition. If v_1, v_2 are eigenvectors for T with possibly different eigenvalues, then $\text{span}\{v_1, v_2\}$ is T-invariant.

More generally, if $v_1, \dots, v_k$ are eigenvectors for T, then $\text{span}\{v_1, \dots, v_k\}$ is T-invariant.

Definition (T-cyclic subspace)**.** The T-cyclic subspace at a vector $v \in V$ is defined as

$$\text{span}\{v, T(v), T^2(v), \dots\} = \text{span}\{T^j(v) : j \in \mathbb{Z}_{\geq 0}\}.$$

Remark (Infinite Span)**.** Recall that if S is a possibly infinite set of vectors in a vector space W, then

$$\text{span}(S) = \left\{ \sum_{j \in S} \alpha_j s_j : \alpha_j \in \mathbb{R}, \text{ and only finitely many } \alpha_j \neq 0 \right\}.$$

This allows us to pick or combine finitely many vectors from S in linear combinations.

Theorem (5.21)**.** Let T be a linear operator on a finite-dimensional vector space V, and let W denote the T-cyclic subspace of V generated by a nonzero vector $v \in V$. Let $k = \dim(W)$. Then:

(a) $\{v, T(v), T^2(v), \dots, T^{k-1}(v)\}$ is a basis for W.

(b) If $a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$, then the characteristic polynomial of $T|_W$ is

$$f(t) = (-1)^k \left(a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k\right).$$

Theorem (5.20)**.** Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Then the characteristic polynomial of $T|_W$ divides the characteristic polynomial of T.

Theorem (Characteristic Polynomial of a Cyclic Subspace)**.** Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V, and let $W \subseteq V$ be the T-cyclic subspace generated by a vector $v \in V$. If $\{v, T(v), T^2(v), \dots, T^{n-1}(v)\}$ is a basis for W, then:

1. The matrix representation of $T|_W$ with respect to this basis is

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix},$$

   where $T^n(v) = -a_0 v - a_1 T(v) - \cdots - a_{n-1} T^{n-1}(v)$.

2. The characteristic polynomial of $T|_W$ is

$$f_{T|W}(t) = (-1)^n (a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + t^n).$$

Proposition (Characteristic Polynomial Decomposition)**.** Let $W \subseteq V$ be a T-invariant subspace of a vector space V. Then the characteristic polynomial of T satisfies the relation

$$f_T(t) = p(t) \cdot q(t),$$

where $p(t)$ is the characteristic polynomial of $T|_W$, and $q(t)$ is the characteristic polynomial of $T|_{V/W}$.

Remark. To see this explicitly, take any basis $\mathcal{B}_1$ for W and extend it to a basis $\mathcal{B}_2 = \mathcal{B}_1 \cup \mathcal{Q}$ for the entire vector space V. Then, the matrix representation of T with respect to $\mathcal{B}_2$ is block-upper triangular:

$$[T]_{\mathcal{B}_2} = \begin{pmatrix} [T|_W]_{\mathcal{B}_1} & A_1 \\ 0 & A_2 \end{pmatrix},$$

where $A_1 = 0$ if and only if $\operatorname{span}(\mathcal{Q})$ is T-invariant. The determinant of $tI_V - [T]_{\mathcal{B}_2}$ decomposes as

$$\det(tI_V - [T]_{\mathcal{B}_2}) = \det(tI_W - [T|_W]_{\mathcal{B}_1}) \cdot \det(tI_{V/W} - A_2),$$

which corresponds to the factorization $f_T(t) = p(t) \cdot q(t)$.

Proposition. If V is T cyclic, then S commutes with T if and only if $S = g(T)$ for polynomial g.

Proof. Assume that V is a cyclic T-module, generated by a vector v, so that

$$V = \operatorname{span}\{v, Tv, T^2v, \dots\}.$$

Let $m(x)$ be the minimal polynomial of T with respect to v, i.e., the monic polynomial of smallest degree such that

$$m(T)v = 0.$$

Then every vector in V can be expressed as a polynomial in T of degree less than $\deg m$ applied to v.

($\Longrightarrow$) Suppose that S commutes with T, i.e., $ST = TS$.

Since V is generated by v, the action of S is determined by its action on v. Let us express Sv as

$$Sv = p(T)v,$$

for some polynomial $p(x)$.

We need to show that $S = p(T)$. For any non-negative integer k,

$$ST^kv = T^kSv = T^kp(T)v = p(T)T^kv.$$

On the other hand,

$$ST^kv = p(T)T^kv.$$

This equality holds for all k, and since $\{v, Tv, T^2v, \dots\}$ spans V, it follows that

$$S = p(T).$$

Therefore, S is a polynomial in T.

($\Longleftarrow$) Conversely, suppose that $S = g(T)$ for some polynomial $g(x)$.

Since polynomials in T commute with T, we have

$$ST = g(T)T = Tg(T) = TS.$$

Thus, S commutes with T.

Combining both directions, we conclude that S commutes with T if and only if $S = g(T)$ for some polynomial g. $\square$

3 Inner Product Spaces and Adjoints

Definition (Standard Inner Product (Real))**.** Let $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$. The standard inner product (dot product) is defined as:

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}.$$

Remark. For $\mathbf{x} \in \mathbb{R}^n$, $\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + \cdots + x_n^2 = \|\mathbf{x}\|^2$.

Definition (Standard Inner Product (Complex))**.** Let $\mathbf{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$. The standard inner product of vectors is defined as:

$$\langle \mathbf{z}, \mathbf{w} \rangle = \overline{z_1} w_1 + \cdots + \overline{z_n} w_n.$$

Remark. For any $\mathbf{w} \in \mathbb{C}^n$, $\langle \mathbf{w}, \mathbf{w} \rangle \in \mathbb{R}_{\geq 0}$, and it is equal to zero if and only if $\mathbf{w} = \mathbf{0}$.

Remark. In $\mathbb{R}^2$, the cosine of the angle θ between two vectors $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is given by:

$$\cos\theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

Definition (Inner Product)**.** An *inner product* on an F-vector space V is the data of a scalar $\langle v, w \rangle \in F$ for every $v, w \in V$, such that the following properties hold:

1. **Linearity in the First Variable:** For all $v_1, v_2, w \in V$ and $\alpha_1, \alpha_2 \in F$,

   $$\langle \alpha_1 v_1 + \alpha_2 v_2, w \rangle = \alpha_1 \langle v_1, w \rangle + \alpha_2 \langle v_2, w \rangle.$$

2. **Conjugate Symmetry:** For all $v, w \in V$,

   $$\langle v, w \rangle = \overline{\langle w, v \rangle}.$$

3. **Positive Definiteness:** If $v \in V$ is a nonzero vector, then

   $$\langle v, v \rangle > 0,$$

   where the result is a positive real number (even if $F = \mathbb{C}$).

The inner product is a map $\langle \cdot, \cdot \rangle : V \times V \to F$.

Definition (Inner Product Space)**.** An *inner product space* is the data of a vector space V over F and an inner product on V.

Corollary (Orthogonal Basis Expansion)**.** Assume V is an inner product space (IPS), and let $\{v_1, \ldots, v_d\}$ be an orthonormal basis (ONB) for V. Then, for any $v \in V$, we have:

$$v = \langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + \cdots + \langle v, v_d \rangle v_d.$$

Theorem (Gram-Schmidt Process)**.** Let $S = \{v_1, \ldots, v_m\}$ be a set of a finite number of vectors in an inner product space (IPS) V. Then, there exists an orthonormal set of vectors $\{v_{r+1}, \ldots, v_d\} \subset V$ such that $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_d\}$ forms an orthonormal basis (ONB) for V.

Definition (Orthogonal Complement)**.** Given a subspace W of an inner product space V, its *orthogonal complement* is defined as:

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

Theorem. If W is any subspace of a finite-dimensional inner product space (IPS) V, then:

$$V = W \oplus W^\perp,$$

where $W^\perp$ is the orthogonal complement of W.

Proof (Sketch). Use the Gram-Schmidt process to construct an orthonormal basis $\{w_1, w_2, \ldots, w_r\}$ for W. Then, extend this basis to an orthonormal basis for V by adding vectors from $W^\perp$. The resulting basis $\{w_1, \ldots, w_r, w_{r+1}, \ldots, w_d\}$ satisfies the decomposition $V = W \oplus W^\perp$. □

Theorem. Fix an inner product space V. The function $P : V \to V^*$, defined by:

$$P(v)(w) = \langle w, v \rangle \quad \text{for } v, w \in V,$$

is a bijection. However, P is not linear over $\mathbb{C}$ if V is a complex vector space.

Theorem. Let $T : V \to V$ be a linear endomorphism of a finite-dimensional inner product space (IPS) V. Then, there exists a unique linear map $T^* : V \to V$ such that:

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle \quad \text{for all } v, w \in V.$$

This function T^* is linear.

Definition (Adjoint or Conjugate Transpose)**.** The linear operator $T^* : V \to V$ is called the *conjugate transpose* or *adjoint* of T.

Theorem. If we choose a basis $\mathcal{B} = \{v_1, \ldots, v_d\}$ for a finite-dimensional inner product space (IPS) V, then the conjugate transpose of T satisfies:

$$[T^*]_\mathcal{B} = ([T]_\mathcal{B})^\dagger,$$

where $([T]_\mathcal{B})^\dagger = ([T]_\mathcal{B})^T$ is the transpose (or conjugate transpose in the complex case) of the matrix representation of T in the basis $\mathcal{B}$.

Theorem. Let $T, U : V \to V$ be linear operators on a finite-dimensional inner product space (IPS) V. Then, the following properties hold:

1. $(U + T)^* = U^* + T^*$,
2. If $\alpha \in F$, then $(\alpha T)^* = \overline{\alpha} T^*$,
3. $(U \circ T)^* = T^* \circ U^*$,
4. $(T^*)^* = T$,
5. $I^* = I$, where I is the identity operator.

Remark. These properties hold because for any composition of operators, $(AB)^* = B^* A^*$, which can be verified using the definition of the adjoint:

$$\langle (AB)v, w \rangle = \langle v, (AB)^* w \rangle = \langle v, B^* A^* w \rangle.$$

4 Spectral Theorem

Definition (Normal Operator)**.** A linear operator $T : V \to V$ on a finite-dimensional inner product space (IPS) V is called *normal* if:

$$TT^* = T^* T.$$

Theorem. If $\mathcal{B}$ is an orthonormal basis of V, then the isomorphism $C_{\mathcal{B}} : V \to \mathbb{F}^d$ has the property that the inner product $\langle v, w \rangle$ is the standard inner product $C_{\mathcal{B}} v \cdot C_{\mathcal{B}} w$.

Theorem. If $\mathcal{B}$ is an orthonormal basis for V, then the matrix representation of T^* satisfies:

$$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^T \quad \text{if } \mathbb{F} = \mathbb{R},$$

and

$$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^H \quad \text{if } \mathbb{F} = \mathbb{C},$$

where $[T]_{\mathcal{B}}^H$ is the conjugate transpose of $[T]_{\mathcal{B}}$.

Proof. Let $\mathcal{B} = \{v_1, \ldots, v_d\}$ be the orthonormal basis. For any $j \in \{1, 2, \ldots, d\}$,

$$\sum_{i=1}^{d} ([T^*]_{\mathcal{B}})_{ij} v_i = T^*(v_j).$$

By the definition of the matrix representation,

$$T^*(v_j) = \sum_{i=1}^{d} \langle T^*(v_j), v_i \rangle v_i.$$

Using the definition of the adjoint,

$$\sum_{i=1}^{d} \langle v_i, T^*(v_j) \rangle v_i = \sum_{i=1}^{d} \langle T(v_i), v_j \rangle v_i.$$

Since $\langle T(v_i), v_j\rangle = ([T]_{\mathcal{B}})_{ji}$, we obtain

$$\sum_{i=1}^{d}([T]_{\mathcal{B}})_{ji}v_i = \sum_{i=1}^{d}([T]_{\mathcal{B}}^{T})_{ij}v_i.$$

Thus, $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^T$ if $\mathbb{F} = \mathbb{R}$ and $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^H$ if $\mathbb{F} = \mathbb{C}$. □

Theorem. Assume $\mathbb{F} = \mathbb{C}$. Then T is normal if and only if there exists an orthonormal basis (ONB) of eigenvectors of T.

(This relies on $\mathbb{F} = \mathbb{C}$ because polynomials split.)

Definition. We say that T is **self-adjoint** if $T = T^*$.

Theorem. If $\mathbb{F} = \mathbb{R}$, then T is self-adjoint if and only if there exists an ONB of eigenvectors.

Remark. Self-adjoint $\Rightarrow$ Normal.

Theorem. Assume T is normal. Then we have:

(a) $\|T(v)\| = \|T^*(v)\|$ for all $v \in V$.

(b) $T - cI$ is normal for all $c \in \mathbb{F}$.

(c) If $v \in V$ is an eigenvector for T with eigenvalue λ, then v is also an eigenvector for T^* with eigenvalue $\overline{\lambda}$.

(d) If v_1, v_2 are eigenvectors for T with distinct eigenvalues, then $\langle v_1, v_2\rangle = 0$.

Theorem. Let T be a normal operator on a finite-dimensional complex inner product space V, and let W be a subspace of V. If W is T-invariant, then W is also T^*-invariant.

Lemma (Normality and polynomial existence)**.** T is normal if and only if there exists a polynomial p such that $T^* = p(T)$.

The forward direction relies on the spectral theorem and polynomial interpolation, i.e., $p(\lambda_i) = \bar{\lambda}_i$

Theorem (6.14)**.** Let S be a linear endomorphism of a finite-dimensional vector space W over an arbitrary field k such that the characteristic polynomial of S splits over k. Then, there exists a basis β of W such that the matrix $[S]_\beta$ is upper triangular.

Moreover, due to Schur's Theorem, if $T : V \to V$ is a linear operator such that its characteristic polynomial splits, then there exists an orthonormal basis β for V such that $[T]_\beta$ is upper triangular.

Theorem (6.16)**.** Assume $F = \mathbb{C}$. Then T has an orthonormal basis of eigenvectors if and only if T is normal.

Theorem (6.17)**.** Assume $F = \mathbb{R}$. Then T has an orthonormal basis of eigenvectors if and only if T is self-adjoint.

Definition. Assume T is a linear endomorphism on an inner product space V. Assume T preserves all lengths, i.e., for all $v \in V$,

$$\langle v, v \rangle = \langle T(v), T(v) \rangle.$$

If our ground field is $\mathbb{R}$, we say T is **orthogonal**. If $F = \mathbb{C}$, we say that T is **unitary**.

Proposition. Any orthogonal or unitary transformation is always invertible if V is finite-dimensional.

Proof. Suppose $v \in \ker(T)$ for an orthogonal (or unitary) operator T. Then:

$$0 = \|T(v)\| = \|v\| \implies v = 0.$$

This implies T is injective. Since V is finite-dimensional, injectivity implies surjectivity, so T is invertible. □

Theorem (Theorem 6.18)**.** Assume $T : V \to V$ is a linear endomorphism of a finite-dimensional inner product space (fdIPS). The following conditions are equivalent:

(a) $T^*T = I$.

(b) $TT^* = I$.

(c) $\langle T(v), T(w) \rangle = \langle v, w \rangle$ for any $v, w \in V$.

(d) If $\beta = \{u_1, \ldots, u_d\}$ is an orthonormal basis (ONB) of V, then $T(\beta) = \{T(u_1), \ldots, T(u_d)\}$ is an ONB of V.

(e) There exists an ONB β for V such that $T(\beta)$ is an ONB for V.

(f) T is *orthogonal* if $\mathbb{F} = \mathbb{R}$ or *unitary* if $\mathbb{F} = \mathbb{C}$.

Lemma. Let $U : V \to V$ be a skew-self-adjoint linear operator on a finite-dimensional inner product space (fdIPS) V such that:

$$\langle v, U(v) \rangle = 0 \quad \text{for all } v \in V.$$

Then $U = 0$.

Proof. Assume $v \in V$. We want to show that $U(v) = 0$.

Consider:

$$0 = \langle v + U(v), U(v + U(v)) \rangle.$$

Expanding the inner product:

$$\langle v, U(v) \rangle + \langle v, U^2(v) \rangle + \langle U(v), U(v) \rangle + \langle U(v), U^2(v) \rangle.$$

Using linearity and self-adjoint properties:

$$\langle v, U^2(v) \rangle + \langle U(v), U(v) \rangle = 2\langle U(v), U(v) \rangle.$$

Since the left-hand side is zero, we conclude:

$$U(v) = 0 \quad \text{for all } v \in V.$$

Thus, $U = 0$ as a linear operator.

As a consequence, if $U = I - T^*T$, we obtain:

$$I = T^*T.$$

□

Corollary. Let $T : V \to V$ be a linear endomorphism of a finite-dimensional inner product space (fdIPS) over $\mathbb{R}$. Then T is self-adjoint and orthogonal if and only if there exists an orthonormal basis (ONB) of eigenvectors

$$\{v_1, \ldots, v_d\}$$

of V such that:

$$T(v_i) = \pm 1 \quad \text{for all } i \in \{1, \ldots, d\}.$$

Definition (Projection onto a Subspace)**.** Let V be a vector space, and let W_1, W_2 be subspaces of V such that:

$$V = W_1 \oplus W_2.$$

The *projection* of V onto W_1, along W_2, is the linear map $\text{proj}_{W_1} : V \to W_1$ given by:

$$\text{proj}_{W_1}(v) = w_1,$$

where $v = w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$.

Note: The kernel and image of this projection satisfy:

$$\ker(\text{proj}_{W_1}) = W_2, \quad \text{Im}(\text{proj}_{W_1}) = W_1.$$

Terminology: We will sometimes simply say that a linear map $T : V \to V$ is a *projection* if it is a projection onto its image along $\ker(T)$.

Proposition. If V is a finite-dimensional inner product space (fdIPS) and $W \subseteq V$ is a subspace, then:

$$V = W \oplus W^\perp.$$

Proof. Assume $\{u_1, \ldots, u_k\}$ is an orthonormal basis (ONB) for W. Extend this to an ONB for V by adding vectors $\{u_{k+1}, \ldots, u_n\}$.

For any $v \in V$, there exist scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that:

$$v = \alpha_1 u_1 + \cdots + \alpha_k u_k + (\alpha_{k+1} u_{k+1} + \cdots + \alpha_n u_n).$$

The first sum belongs to W, and the second sum belongs to $W^\perp$.

We claim:

$$\alpha_1 u_1 + \cdots + \alpha_k u_k \in W, \quad \alpha_{k+1} u_{k+1} + \cdots + \alpha_n u_n \in W^\perp.$$

Moreover, this decomposition is unique, ensuring that:

$$V = W \oplus W^\perp.$$

□

Definition (Orthogonal Projection)**.** If V is a finite-dimensional inner product space (fdIPS) and $W \subseteq V$ is a subspace, the *orthogonal projection* onto W is defined as the projection onto W along $W^\perp$.

Theorem (Spectral Theorem, Theorem 6.25)**.** Assume $T : V \to V$ is a linear endomorphism of a finite-dimensional inner product space (fdIPS) V that is self-adjoint. If $\mathbb{F} = \mathbb{R}$ or if T is normal when $\mathbb{F} = \mathbb{C}$, then:

Let $W_1, \ldots, W_r$ denote the eigenspaces corresponding to the distinct eigenvalues $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$.

Furthermore, for each $i \in \{1, \ldots, r\}$, let $T_i : V \to V$ denote the orthogonal projection onto W_i. Then the following hold:

(i) $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$.

(ii) For each fixed $i \in \{1, \ldots, r\}$,

$$W_i = \left(\bigoplus_{j \neq i} W_j \right)^{\perp}.$$

(iii) For all $i, j \in \{1, \ldots, r\}$,

$$T_i T_j = \delta_{ij} T_j.$$

(iv) The operator T decomposes as:

$$T = \lambda_1 T_1 + \cdots + \lambda_r T_r.$$

(v) The identity operator decomposes as:

$$I = T_1 + \cdots + T_r.$$

Definition (Spectrum)**.** The set of eigenvalues of an operator T, as given in the spectral theorem, is called the *spectrum* of T.

Corollary. A linear endomorphism $T : V \to V$ of a finite-dimensional inner product space (fdIPS) V is *unitary* if and only if T is normal and all eigenvalues of T have length 1.

Remark. If $z \in \mathbb{C}$ and $z = a + bi$, then the inverse of z is given by:

$$z^{-1} = \frac{a - bi}{a^2 + b^2}.$$

In particular, if $|z| = 1$, then:

$$z^{-1} = \overline{z}.$$

Corollary. Assume V is a finite-dimensional inner product space (fdIPS), and $T : V \to V$ is a linear endomorphism of V. Then, T is self-adjoint if and only if T is normal and all eigenvalues of T are real.

5 Geometry of Orthogonal Operators

Theorem (Theorem 6.23: Orthogonal Real Transformations in two dimensions)**.** Assume $T : \mathbb{R}^2 \to \mathbb{R}^2$ is an orthogonal linear transformation. What are all possible such linear transformations?

1. The determinant of T satisfies:

$$\det(T) = \pm 1.$$

2. If $\det(T) = 1$, then T is a *rotation*, and can be expressed as:

$$T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

for some angle θ.

Definition (Reflection)**.** The *reflection* about the line span $\left\{\begin{bmatrix}0\\1\end{bmatrix}\right\}$ is the function:

$$R_{\text{span}\{(0,1)\}} : \mathbb{R}^2 \to \mathbb{R}^2$$

given by the transformation matrix:

$$R_{\text{span}\{(0,1)\}} = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}.$$

Definition (Reflection Across a Line)**.** Reflection across the line

$$\text{span}\left\{\begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}\right\}$$

for some $\theta \in \mathbb{R}$ is the function $R : \mathbb{R}^2 \to \mathbb{R}^2$ given by left multiplication by:

$$R = \begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{bmatrix}.$$

Expanding the computation:

$$\begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix} = \begin{bmatrix}\cos\theta & \sin\theta\\\sin\theta & -\cos\theta\end{bmatrix}.$$

Multiplying with the final matrix:

$$\begin{bmatrix}\cos\theta & \sin\theta\\\sin\theta & -\cos\theta\end{bmatrix}\begin{bmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{bmatrix} = \begin{bmatrix}\cos^2\theta - \sin^2\theta & 2\sin\theta\cos\theta\\2\sin\theta\cos\theta & \sin^2\theta - \cos^2\theta\end{bmatrix}.$$

Using trigonometric identities, we simplify:

$$\begin{bmatrix}\cos 2\theta & \sin 2\theta\\\sin 2\theta & -\cos 2\theta\end{bmatrix}.$$

Thus, the reflection matrix is:

$$R = \begin{bmatrix}\cos 2\theta & \sin 2\theta\\\sin 2\theta & -\cos 2\theta\end{bmatrix}.$$

Theorem (Theorem 6.23)**.** Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be an orthogonal linear endomorphism of $\mathbb{R}^2$ with its standard inner product. Then, exactly one of the following holds:

1. T is a *rotation* by some angle $\theta \in [0, 2\pi]$, and $\det(T) = 1$.

2. T is a *reflection* about some line passing through the origin, and $\det(T) = -1$.

Definition (Rotations and Reflections in a Subspace)**.** Let W be a two-dimensional subspace of an inner product space V. We say that $T : W \to W$ is a *rotation* if there exists some orthonormal basis $\beta = \{u_1, u_2\}$ such that the matrix representation of T in this basis is:

$$[T]_\beta = \begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}$$

for some $\theta \in [0, 2\pi]$.

If $T : W \to W$ is a linear endomorphism and there exists an orthonormal basis of W such that:

$$[T]_\beta = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

then we say that T is a *reflection*.

If W is a one-dimensional subspace and $T : W \to W$ is the function satisfying $T(w) = -w$ for all $w \in W$, we say that T is a *reflection*. On the other hand, if $T(w) = w$ for all $w \in W$, we say that T is a *rotation*.

Corollary (Corollary 6.46)**.** The composite of a reflection and a rotation is a reflection.

Lemma. If $T : V \to V$ is an orthogonal endomorphism on a finite-dimensional inner product space (fdIPS) V, then there exists some subspace W such that:

$$1 \le \dim(W) \le 2$$

and W is T-invariant if $\dim(V) \ge 1$.

Theorem. Assume $T : V \to V$ is an orthogonal operator on a finite-dimensional nonzero real inner product space (RIPS). Then, there exist mutually orthogonal subspaces $W_1, W_2, \ldots, W_m$ such that all the following hold:

(i) $\dim(W_i) \in \{1, 2\}$ for all $i \in \{1, \ldots, m\}$.

(ii) W_i is T-invariant for all $i \in \{1, \ldots, m\}$.

(iii) (Knowing $T|_{W_i} : W_i \to W_i$) we have the decomposition:

$$V = \bigoplus_{i=1}^{m} W_i.$$

Theorem. Moreover:

(A) The number of subspaces W_i for which $T|_{W_i}$ is a reflection (as opposed to a rotation) is even if $\det(T) = 1$ and odd if $\det(T) = -1$.

(B) In *fact*, there exists a decomposition $W_1, \ldots, W_m \subseteq V$ satisfying (i)–(iii) such that, for at most one $i \in \{1, \ldots, m\}$, $T|_{W_i}$ is a reflection.

Example. Using Theorem 6.47, set up a look at the case where $\dim(V) = 3$. One of the following holds:

Case 1: If $n = 3$, then $\dim(W_i) = 1$ for $i \in \{1, 2, 3\}$. So the matrix of T is

$$\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}$$

with respect to some orthonormal basis.

Case 2: If $n = 2$, then there exists an orthonormal basis such that the restriction of T has the matrix

$$\begin{pmatrix} R & 0 \\ 0 & \pm 1 \end{pmatrix}$$

where R is either a reflection or a rotation.

6 Bilinear Forms

Definition. A **bilinear form** on a k-vector space V is a function

$$H : V \times V \to k$$

such that for any $v_1, v_2, w \in V$ and scalars $\alpha, \beta \in k$, the following hold:

(a) $H(\alpha v_1 + \beta v_2, w) = \alpha H(v_1, w) + \beta H(v_2, w)$

(b) $H(w, \alpha v_1 + \beta v_2) = \alpha H(w, v_1) + \beta H(w, v_2)$

Example. If $A \in k^{d\times d}$, we can define a bilinear form

$$H : k^d \times k^d \to k$$

by the formula

$$H(v, w) = v^\top Aw.$$

(Proof left as an exercise.)

Definition. Given two bilinear forms on a vector space V, say H_1, H_2, their sum is the function

$$(H_1 + H_2) : V \times V \to k$$

such that

$$(H_1 + H_2)(v, w) = H_1(v, w) + H_2(v, w), \quad \forall v, w \in V.$$

If H is any bilinear form and $\alpha \in k$, the scalar product $\alpha H : V \times V \to k$ is the function

$$(\alpha H)(v, w) = \alpha(H(v, w)).$$

Theorem (6.31 / Exercise)**.** The sum of any two bilinear forms is a bilinear form, and the scalar product of any scalar and any bilinear form is a bilinear form. Moreover, if $B(V)$ is the set of bilinear forms on a vector space, then $B(V)$ has a vector space structure over k.

Definition. Assume V is a finite-dimensional vector space with basis

$$\mathcal{B} = \{v_1, \ldots, v_d\}.$$

Then, the matrix representation with respect to $\mathcal{B}$ of a bilinear form $H : V \times V \to k$ is the matrix whose (i, j) entry is given by

$$H_{ij} = H(v_i, v_j), \quad \forall i, j \in \{1, \ldots, d\}.$$

Theorem. For every basis $\mathcal{B}$ of a finite-dimensional vector space V and every bilinear form $H : V \times V \to k$, we use the notation

$$T_{\mathcal{B}}(H) \in k^{d\times d}$$

to denote the matrix representation of H.

Theorem (Theorem 6.32)**.** If V is a finite-dimensional vector space and

$$\mathcal{B} = \{v_1, \dots, v_d\}$$

is a basis for V, then

$$T_{\mathcal{B}} : \mathbb{B}(V) \to k^{d \times d}$$

is an isomorphism of vector spaces.

Corollary (2-3)**.** Let V be a finite-dimensional vector space, and let H be a bilinear form on V. If $\mathcal{B} = \{v_1, \dots, v_d\}$ is a basis for V, then

$$H(v, w) = [v]_{\mathcal{B}}^T \Psi_{\mathcal{B}}(H)[w]_{\mathcal{B}},$$

where $\Psi_{\mathcal{B}}(H)$ is the matrix representation of H in the basis $\mathcal{B}$.

In particular, if $V = k^d$ and $\mathcal{B}$ is the standard basis, then any bilinear form $H : k^d \times k^d \to k$ has the property that

$$H(x, y) = x^T \operatorname{Tab}(H) y.$$

Theorem (6.33)**.** Let V be a finite-dimensional vector space, and let $\mathcal{B}, \mathcal{B}'$ be two bases for V. If H is a bilinear form on V, then its matrix representation changes as follows:

$$\Psi_{\mathcal{B}'}(H) = I_{\mathcal{B}' \to \mathcal{B}}^T \Psi_{\mathcal{B}}(H) I_{\mathcal{B}' \to \mathcal{B}}.$$

Proof. Let $\mathcal{B} = \{v_1, \dots, v_d\}$ and let $\mathcal{B}' = \{w_1, \dots, w_d\}$. Then, for any $w_i, w_j \in \mathcal{B}'$, we express the bilinear form as:

$$H(w_i, w_j) = [w_i]_{\mathcal{B}'}^T \Psi_{\mathcal{B}'}(H)[w_j]_{\mathcal{B}'}.$$

Since the basis transformation satisfies $[w_i]_{\mathcal{B}'} = I_{\mathcal{B}' \to \mathcal{B}}[w_i]_{\mathcal{B}}$, we substitute:

$$H(w_i, w_j) = (I_{\mathcal{B}' \to \mathcal{B}}[w_i]_{\mathcal{B}})^T \Psi_{\mathcal{B}'}(H)(I_{\mathcal{B}' \to \mathcal{B}}[w_j]_{\mathcal{B}}).$$

Rewriting,

$$H(w_i, w_j) = [w_i]_{\mathcal{B}}^T I_{\mathcal{B}' \to \mathcal{B}}^T \Psi_{\mathcal{B}'}(H) I_{\mathcal{B}' \to \mathcal{B}}[w_j]_{\mathcal{B}}.$$

Since this holds for all $w_i, w_j \in V$, we conclude:

$$\Psi_{\mathcal{B}'}(H) = I_{\mathcal{B}' \to \mathcal{B}}^T \Psi_{\mathcal{B}}(H) I_{\mathcal{B}' \to \mathcal{B}}.$$

□

Definition. Let $P, Q \in k^{d \times d}$ for $d \geq 2$, or more generally for $d \geq 0$ in an infinite-dimensional setting. We say P and Q are **congruent** if there exists an invertible matrix M such that

$$P = M^T Q M.$$

Definition. A bilinear form $H : V \times V \to k$ is **symmetric** if

$$H(v, w) = H(w, v) \quad \text{for all } v, w \in V.$$

Theorem (6.34)**.** Let H be a bilinear form on a finite-dimensional vector space V, and let β be an ordered basis for V. Then H is symmetric if and only if $\psi_\beta(H)$ is symmetric.

Lemma. Let H be a nonzero symmetric bilinear form on a vector space V over a field F not of characteristic two. Then there exists a vector $x \in V$ such that $H(x, x) \neq 0$.

Theorem (Theorem 6.35)**.** Let V be a finite-dimensional vector space over a field F not of characteristic two. Then every symmetric bilinear form on V is diagonalizable.

Proof. We use mathematical induction on $n = \dim(V)$. If $n = 1$, then every element of $B(V)$ is diagonalizable.

Now suppose that the theorem is valid for vector spaces of dimension less than n for some fixed integer $n > 1$, and suppose that $\dim(V) = n$. If H is the zero bilinear form on V, then trivially H is diagonalizable; so suppose that H is a nonzero symmetric bilinear form on V.

By the lemma, there exists a nonzero vector $x \in V$ such that $H(x, x) \neq 0$. Recall the function $L_x : V \to F$ defined by

$$L_x(y) = H(x, y) \quad \text{for all } y \in V.$$

By a standard property of bilinear forms, L_x is linear. Furthermore, since $L_x(x) = H(x, x) \neq 0$, we have that L_x is nonzero. Consequently, $\operatorname{rank}(L_x) = 1$, and hence $\dim(N(L_x)) = n - 1$.

The restriction of H to $N(L_x)$ is obviously a symmetric bilinear form on a vector space of dimension $n - 1$. Thus, by the induction hypothesis, there exists an ordered basis $\{v_1, v_2, \ldots, v_{n-1}\}$ for $N(L_x)$ such that

$$H(v_i, v_j) = 0 \quad \text{for } i \neq j, \quad (1 \leq i, j \leq n - 1).$$

Set $v_n = x$. Then $v_n \notin N(L_x)$, and so $\beta = \{v_1, v_2, \ldots, v_n\}$ is an ordered basis for V. In addition,

$$H(v_i, v_n) = H(v_n, v_i) = 0 \quad \text{for } i = 1, 2, \ldots, n - 1.$$

We conclude that $\psi_\beta(H)$ is a diagonal matrix, and therefore H is diagonalizable. □

Corollary. Let F be a field that is not of characteristic two. If $A \in M_{n\times n}(F)$ is a symmetric matrix, then A is congruent to a diagonal matrix.

Proposition. Let $k = \mathbb{Z}/2\mathbb{Z}$. The function

$$H : k^2 \times k^2 \to k$$

is given by

$$H\left((x_1, x_2), (y_1, y_2)\right) = x_1 y_2 + x_2 y_1.$$

I claim H is symmetric but not diagonalizable.

(H is symmetric)

Definition. The **rank** of a bilinear form

$$H : V \times V \to k$$

on a finite-dimensional vector space V is the rank of $\psi_\beta(H)$ for any basis β of V.